(12) UK Patent Application (19) GB (11) 2 295 152 (13) A

(43) Date of A Publication 22.05.1996

(21) Application No 9423332.7

(22) Date of Filing 18.11.1994

(71) Applicant(s)
Pfizer Limited

(Incorporated in the United Kingdom)

Patent Department, Ramsgate Road, SANDWICH, Kent, CT13 9NJ, United Kingdom

(72) Inventor(s)
Ryszard Jurek Kobylecki
John Mark Francis Gardner

(74) Agent and/or Address for Service
Urquhart-Dykes & Lord
91 Wimpole Street, LONDON, W1M 8AH, United Kingdom

(51) INT CL[6]
C07K 1/04 , B32B 5/26 7/12 , D21H 17/26

(52) UK CL (Edition O)
C3H HA3 H303 H370 H380
B5N N0526 N0712 N175 N176 N177 N178 N18X N207 N21X N22Y N224 N256 N26X N26Y N263 N2702 N2732 N401 N418 N420 N46X N476 N494 N566 N611 N648 N658 N66Y N661 N670 N672
D1R REX R151 R162 R318 R461 R513 R614 R658
D2W WA W326 W401 W412
U1S S1347 S1524 S1537 S1577 S1586 S1591

(56) Documents Cited
WO 93/06121 A1 WO 92/10092 A1
Bioorg.Med.Chem.Lett. 1993,3(3),425-430 Peptide Research 1994,7(6),292-298 Drug Dev.Res. 1994, 33(2), 174-188

(58) Field of Search
UK CL (Edition N) C3H HA3 HB4A HB4B
INT CL[6] C07H 21/00 , C07K 1/04
ONLINE DATABASES: WPI,CLAIMS,DIALOG/BIOTECH

(54) Preparation of a library of compounds by solid-phase synthesis

(57) A method of making a library of compounds comprises the following steps:
(a) individually identifying a plurality of discrete reaction zones defined on laminar solid support material;
(b) charging each of said reaction zones with a starting material;
(c) sub-dividing the reaction zones into at least two initial batches;
(d) applying at least two different reagents, one to each of the reaction zones in each initial batch, and recording the identity of those reaction zones to which each of said different reagents is applied;
(e) subjecting all reaction zones to reaction conditions which promote reaction to completion;
(f) further sub-dividing the reaction zones into at least two alternative batches;
(g) applying at least two different reagents, one to each of the reaction zones in each alternative batch, and recording the identity of those reaction zones to which each of said different reagents is applied;
(h) subjecting all reaction zones to reaction conditions which promote reaction to completion, and
(i) repeating steps (f) to (h) inclusive as many times as desired.

The solid support may comprise paper prepared from cellulose comprising an amino group, obtained by reaction of the cellulose either with an amino precursor, followed by generation of free amine which is blocked by a protecting group, or with a compound having a protected amine group, followed by deprotection. The lamina may comprise a functionalised resin (aminomethylpolystyrene) sandwiched (with a polyethylene) between porous inert sheets (non-woven fibrous polypropylene).

GB 2 295 152 A

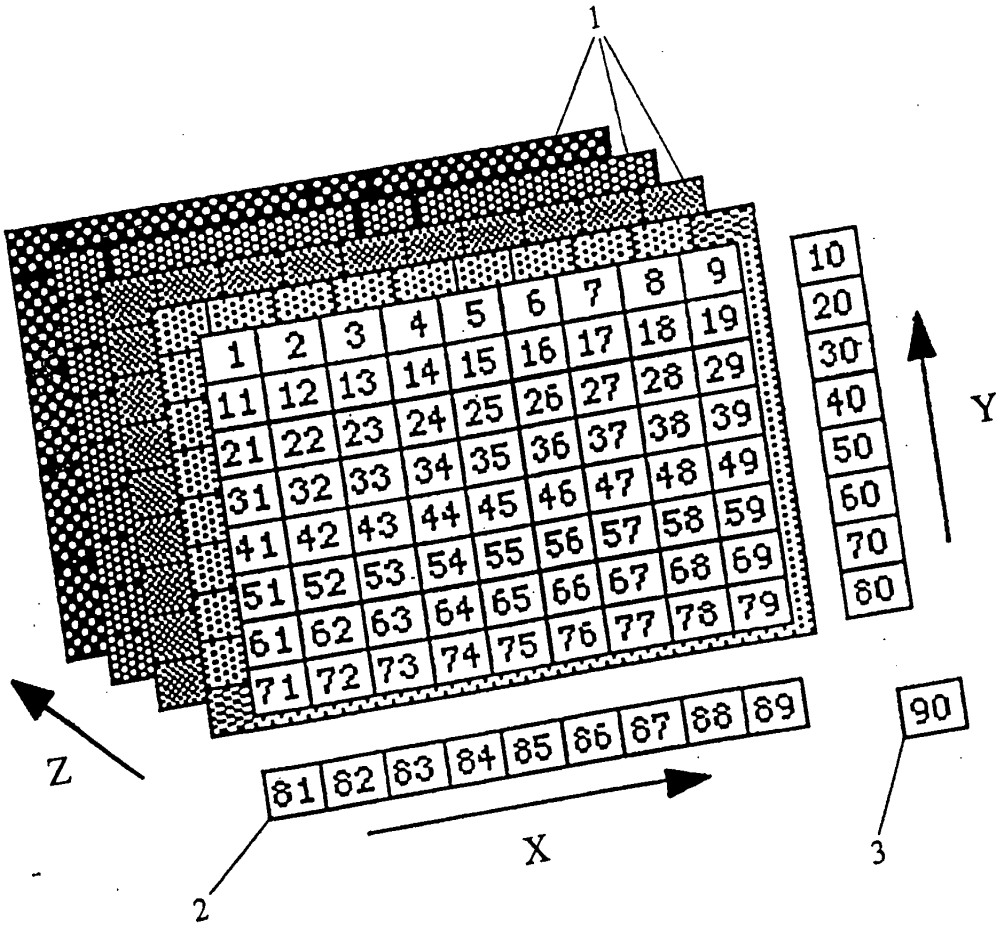
1
1 2 3 4 5 6 7 8 9
11 12 13 14 15 16 17 18 19
21 22 23 24 25 26 27 28 29
31 32 33 34 35 36 37 38 39
41 42 43 44 45 46 47 48 49
51 52 53 54 55 56 57 58 59
61 62 63 64 65 66 67 68 69
71 72 73 74 75 76 77 78 79
10
20
30
40
50
60
70
80
Y
Z
81 82 83 84 85 86 87 88 89
X
2
90
3

FIGURE 1

METHOD OF MAKING A LIBRARY OF COMPOUNDS

The present invention relates to a method of preparation of chemical compounds and, in particular, to a method of preparing combinatorial libraries of chemical compounds. The method is especially suitable for the preparation of natural and synthetic chemical compounds which are to be tested for activity as therapeutic agents, though it need not be used exclusively for this purpose. In addition to being used for the preparation of combinatorial libraries, the method of the present invention also facilitates easy identification of individual compounds, so that any compounds which show encouraging biological activity can be prepared on a larger scale for further analysis. By modifying the method of the present invention, it is possible to prepare individual compounds in pure form in a non-combinatorial format.

The synthesis and screening of combinatorial libraries is becoming increasingly important in the pharmaceutical industry as a means of drug "discovery". The major advantages of combinatorial chemistry are that it is faster and cheaper than orthodox methods. This makes it a much more effective technique in the quest to uncover new therapeutic agents, particularly in circumstances where there is little or no information available concerning the types of structures likely to show the desired activity.

The wider availability of solid-phase synthetic methods has also led to increased interest in combinatorial chemistry. Clearly, solution chemistry is unsuitable for a technique which aims to produce a multiplicity of new products together, since this does not allow physical separation between the different materials produced. The products are therefore likely be contaminated with excess reagents, by-products etc, leading to difficulties in separation and purification.

The preparation of combinatorial compound libraries typically involves a number of successive stages, each of which involves a chemical or enzymatic modification of an existing molecule. Most typically, this process involves the addition of a monomeric unit or other synthon to a growing sequence, or the modification of chemical functionality on the sequence. Conveniently, the sequence or growing chain of interest is attached to a solid support. By carrying out the desired series of synthetic steps on the bound starting material, and by altering the nature of the monomeric or other synthon units, the type of chemistry and the sequence of reactions, it is possible to prepare an enormous number of individual compounds in short time.

As indicated above, combinatorial methods entail a series of chemical steps with multiple choices of chemical reagents for each step. The complexity of the combinatorial library thus produced is determined by the product of the number of reagent choices for each step of the synthesis, which can be quite large. The problem which then arises is identification and characterisation of members of the library which display particular desired properties.

Various solutions have been proposed to deal with this: For example, members of the library can be synthesised in spatially segregated arrays. However, because of the extra burden which maintenance of segregation imposes, this approach tends to lead to relatively small libraries. Alternatively, in the so-called "multivalent synthesis" method, a library of moderate complexity is produced by pooling multiple choices of reagents during synthesis. If a pool is shown to have properties of interest, it is re-synthesised with progressively lower complexity until a single compound or class of compounds is identified having the desired property. The ultimate size of a library produced by this technique is inevitably restricted because of concentration effects which determine the limits of detection at which activity can be discerned.

The so-called "mix and split synthesis" method relies on combinatorial synthesis carried out on discrete solid particles such as minute resin beads. Through a protocol of mixing and separating beads at the end of each step in the synthetic sequence, populations of beads are generated to which are bound the products of specific reaction sequences. Inevitably, individual beads obtained from the final reaction step have different products attached, so that identification and characterisation of active materials is still a problem.

Fortunately, biologically active compounds show remarkable potency and receptor sites are highly selective, so it is possible to detect low concentrations of active compound amid an extensive background of inactive material using standard *in vitro* screening techniques.

Another drawback of the mix and split synthesis method is that some measure of over-representation and omission of individual compounds is inevitable because of the randomness introduced by the mixing and splitting steps.

To counteract the above problems of identification and characterisation, some workers have proposed co-synthesis of a sequencable tag which encodes the series of steps and reagents used during synthesis of respective constituents of the library. More recently, it has been proposed to use tagging molecules to encode both the step number and the chemical reagent used in a given step, as a binary record of the synthetic steps experienced by each bead. This technique undoubtedly adds to the complexity of operations carried out during development of a combinatorial library.

From the foregoing, it is apparent that known methods of preparing combinatorial libraries of chemical compounds suffer from two major drawbacks: Either the materials are prepared by maintaining segregation, with the inevitable consequence that only relatively small libraries are practicable, or the materials are prepared without segregation but in such minute quantities that characterisation is rendered very difficult.

It is therefore an object of the present invention to provide a method of making a library of chemical compounds which allows wide diversification in the products obtained without over-representation and/or omission, at the same time as providing a clear indication of the sequence of steps which has been followed to synthesise a particular compound, thereby facilitating characterisation of individual materials.

In a first aspect, the invention is a method of making a library of compounds, which method comprises the following steps:

(a) individually marking with indicia a plurality of discrete reaction zones defined on laminar solid support material;

(b) charging each of said reaction zones with a starting material;

(c) sub-dividing the reaction zones into at least two initial batches;

(d) applying at least two different reagents, one to each of the reaction zones in each initial batch, and recording the identity of those reaction zones to which each of said different reagents is applied;

(e) subjecting all reaction zones to reaction conditions which promote reaction to completion;

(f) further sub-dividing the reaction zones into at least two alternative batches;

(g) applying at least two different reagents, one to each of the reaction zones in each alternative batch, and recording the identity of those reaction zones to which each of said different reagents is applied;

(h) subjecting all reaction zones to reaction conditions which promote reaction to completion, and

(i) repeating steps (f) to (h) inclusive from zero to n times, as desired.

It will be understood that n may be any whole number integer, the value of which depends on the complexity of the combinatorial library that it is intended to produce.

The method outlined above provides the synthetic chemist for the first time with the means to synthesise any number of single, easily identifiable labelled chemical compounds on a controllable pre-defined scale of preparation. In particular, this invention offers considerable handling advantages over prior art methods. For example, if desired the entire set of individual reaction zones may be handled as a single laminar medium. This opportunity does not exist with free-flowing microscopic resin beads. The method of division does not rely on the laminar support material being a particular shape. Thus, it is possible for the support to be in the form of tapes or streamers.

In an especially preferred form, the reaction zones are defined on sheets of material. An individual sheet may represent a single reaction zone, in which case a plurality of sheets is required to put the invention into effect. Alternatively, a single sheet may be sub-divided into an array of reaction zones of equal size, individual elements of the array being separable from each other for effecting step (c) above. In one possible variant of this method, each sheet is charged with a different starting material in step (b).

An especially preferred form of sheet material is paper, particularly paper which has been treated to enable the starting material to bind to the sheet. When the starting materials are amino acids or peptide fragments, the paper may for example carry allylic anchor groups to releasably bind the carboxylic acid groups of amino acids to the paper; a variety of other linking groups is also possible. The first and subsequent reagents may attach further amino acids or peptide fragments to the already bound amino acid residues on the sheet in known manner.

Another type of paper which may be used has free amino groups which may releasably bind to carboxy groups of amino

acids forming the starting material of the library compounds. One method of making such a paper is to treat cellulose, preferably in powder form, with acrylonitrile and a base, generally under aqueous conditions, to form a cyanoethyl ether of cellulose. The product may be dried and reduced, for example with borane in tetrahydrofuran, to aminopropyl cellulose. After removal of residual reagents the amino groups may be protected, for example by conversion of the aminopropyl groups to *tert*-butyloxycarbonyl aminopropyl groups and the resulting substituted cellulose may be mixed with cellulose fibre and formed into paper by standard paper-making methods.

The *tert*-butyloxycarbonyl or "Boc" groups may then be removed to provide the required paper with free amino groups.

In a second aspect, the invention is a method of preparing a paper support material for use in the synthesis of chemical compound libraries, which method comprises:

(a) linking cellulose with a compound which is selected from the set consisting of an amine precursor or a compound having a protected amine group;

(b) in the case of an amine precursor, generating the free amine and then protecting it with a conventional amino protecting group;

(c) incorporating the amine-functionalised cellulose into a paper sheet by mixing with paper fibre and forming into sheets, and

(d) reacting the paper sheets obtained from step (c) above with an amino deprotecting reagent to provide free amine groups on the paper sheets.

Alternatively, materials other than paper may be used for making the sheets. This is an important consideration for those branches of chemistry which require a non-protic environment, since paper is a protic material.

One possible alternative is a polyethylene or polypropylene film which has been grafted with polystyrene chains, as described in published PCT Patent Application No. WO 90/02749. Alternatively, the sheet may be of a laminated construction, being in the form of a solid material trapped between two or more layers of porous mesh. One laminate of this type consists of a so-called "resin cloth" comprising cross-linked polystyrene resin containing amino groups formed as a layer sandwiched between fibrous sheets, for example, non-woven polypropylene sheets, on which indicia may be borne. The use of other materials is, of course, possible.

A non-protic sandwich material such as that described above permits a wider range of chemistries to be carried out. For example, chemistry is permitted to be performed on a supported resin cloth which usually requires strictly anhydrous conditions. Example reactions include, but are not limited to, use of a strong non-protic base to generate anions of chemical substrates affixed to the resin cloth. Further manipulations of these anions permits, for example, Heck type couplings, Stille couplings, heteroaryl couplings, carbonylations, carboxylations and carbamoylations not normally permitted in a protic environment.

In a third aspect, the invention is a method of preparing a laminar resin support material for use in the synthesis of chemical compound libraries, which method comprises affixing a layer of particulate functionalised solid support resin material to a porous inert laminar material.

Preferably, the layer of particulate functionalised solid support resin material is sandwiched between two layers of porous inert laminar material.

In general, suitable sheet material may be any material which is readily markable with indelible indicia, is divisible in equal proportions, allows the sheets to be formed into a stack and subsequently separated and to which the constituents of the compounds of the library may releasably be bound. The sheet and method of binding the compounds are preferably such

that known amounts of the compound may be repeatably released from a single sheet portion bearing the compound.

It should be noted that the present invention is not limited to the preparation of biologically active compounds. It is applicable to any organic or inorganic species which, used in combination with other reagents, will form oligomers bound to the solid sheet. The compounds which are stored in the library must however be compatible with the material of the sheet.

Besides sheet materials, any suitable support material can be adapted for use in the method of the present invention provided that it has the capacity for division and subsequent sub-division into discrete reaction zones and provided that it possesses the necessary surface active qualities to serve as a vehicle for the intended reaction steps.

The compounds prepared in the library may be linked to the support by a wide variety of methods, depending on the nature of the support and the compounds to be prepared. Apart from the allylic anchoring group/amino acid system mentioned above, chemical linkers which may be cleaved by acidic, basic, hydrogenolytic or other chemical reagents may be used, as may light-induced cleavage. Combinations of these methods may also be used.

The amount of compound stored in each reaction zone may vary according to the nature of the compound and the nature of the support material, and also according to the size of the zones. Amounts of compound varying from a few nanograms to several milligrams may be stored on portions of paper sheet of convenient size. In principle any amount of compound may be stored provided that the support material is large enough.

Typically, the different reagents used in the method according to the present invention are individual monomeric units and may be chosen from a large variety of compounds. These include agents such as amino acids, nucleotides, sugars, naturally-occurring and synthetic heterocycles, lipids, and

combinations thereof, though it will be understood that this list is not exhaustive. In general, any bifunctional group may be used which may be linked to the support material or to the growing sequence in protected form, and subsequently deprotected and reacted with a further group. Alternatively, a monofunctional group may be used to complete the sequence.

It is an essential feature of the present invention that individual reaction zones are identified, that is to say, labelled with some form of indicia which uniquely characterises each reaction zone. The indicia may comprise, for example, numbers, letters, symbols or colours in a coded combination. The indicia may be applied to the respective reaction zones before synthesis commences using known printing methods. These are preferably such that the ink used will not leach out of the reaction zones during the synthetic procedures, or otherwise interfere with formation and subsequent removal of a compound held on a particular reaction zone. U.V.-sensitive ink which is "fixed" to the reaction zones by exposure to ultraviolet radiation after printing is generally suitable for this purpose. Other types of indicia, not necessarily optical in nature, may be used for identifying individual reaction zones. Possible alternatives include Smiles strings, bar-codes, chemical structures, marked or printed punched card formats, ultraviolet-readable fluorescent systems and electro-magnetically readable devices such as magnetic strips. The type of indicia used may depend on the size and shape of the support material and/or reaction zones.

The invention will now be described by way of example only with reference to the drawing (Figure 1) which shows in schematic form one particular embodiment of support material used in performance of the present invention and a convenient pattern of sub-division.

Referring now to Figure 1, the illustrated arrangement shows orthogonal arrays of reaction zones 3 defined on a series of support sheets 1. The reaction zones are arranged in a grid or matrix layout in straight rows along one

dimension and straight columns along the other dimension, each reaction zone being provided with a unique tag or label.

In the next step, each of the sheets 1 is treated with a different first reagent which becomes bound to the sheet to form the first monomer or constituent serving as the starting material for subsequent steps. The sheets are then superposed to form a block in which corresponding reaction zones 3 of respective sheets are aligned with each other. The block of sheets so formed is then divided by making a first series of cuts through the stack, e.g. in the X direction, thereby forming a plurality of stacked strips 2.

Each stack of strips 2 is then treated with a reagent to effect deprotection or activation of the first constituent following reaction with a different second reagent to effect binding of a respective second constituent to the first constituents already bound on the strips.

Following this, the treated stacks of strips are reassembled to reform the block and a second series of cuts is made at right angles (in the Y direction) to the first so that each strip becomes further sub-divided into smaller elements corresponding to the reaction zones (3).

Each of the stacks of individual reaction zones is then deprotected if necessary and treated with a different third reagent to effect binding of a respective third constituent on the free end of the second constituent already in place.

If, in this example, a total of twenty sheets is used initially and if each sheet is treated with a different first reagent monomeric unit, twenty different sheets having attached a first monomer or fragment will be formed. When the superposed sheets are divided to form, say, 20 strips and each strip is treated with a different reagent a total of 20x20=400 dimeric chains is formed, each having a different combination of first and second monomers or fragments. Subsequent reassembly of the block and further sub-division al ng the second dimension into, say, 20 slices and treatment of each of the slices with different reagents will give 20x20x20=8,000

different pure trimeric structures. The total number of monomers, dimers and trimers may be increased by dividing the block into a greater number of strips or slices, or by increasing the number of sheets. Thus, if 50 sheets are used, divided into 50 strips and 50 slices, the number of different pure individual trimeric structures will be 125,000.

Each of the trimers will be different and will be identifiable unambiguously from the indicia (which may be letters and/or numbers applied by printing) marked on the reaction zones.

At this final stage of the process, the sheets have been cut in a fashion to provide individual pieces of paper, each of which is marked with a single, unique index, which is in itself an identifier of the single, unique chemical structure attached to that portion of paper. Furthermore, all possible combinations are formed of compounds available from the constituents provided by the reagents used.

In the embodiment of the invention described above, the reaction zones may be square or oblong and arranged in an orthogonal pattern. However, other geometrical arrangements may be used. In principle the sheet portions may be of any shape and arranged in any type of grid pattern, subject only to the need to divide the sheet into individual portions.

The sheet material may be, but is not limited to, paper and depending on the size of the sheets cutting may be carried out using any suitable cutting device, such as scissors or an ordinary office guillotine. The arrangement described above allows a very large number of different dimeric and trimeric and larger polymeric structures to be assembled easily and rapidly.

It will be apparent that in a library of single individual compounds, each of which is identified by means of its own unique indicia, an individual sheet portion may be easily identified. Thus, evaluation of the biological or other activity of the compound cleaved from such an identified sheet portion will permit, by means of target d screening of